

ABSTRACT

The present invention provides a method and apparatus for retrieving documents that are stored in a language other than the language that is used to formulate a search query. This invention decomposes the query into terms and then translates each of the terms into terms of the language of the database. Once the database language terms have been listed, a series of subqueries is formed by creating all the possible combinations of the listed terms. Each subquery is then scored on each of the documents in the target language database. Only those subqueries that return meaningful scores are relevant to the query. Thus, the semantic meaning of the query is determined against the database itself and those documents in the database language that are most relevant to that semantic meaning are returned.